



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/650,365

Source: HPWO

Date Processed by STIC: 11/21/03

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION AND PATENT IN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221**

**Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail. Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>) , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

IFwo

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/650,365

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos     The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length     The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering     The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII     The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length     Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"     A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)     Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence  
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
 This sequence is intentionally skipped  
  
 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences
- 8      Skipped Sequences  
    (NEW RULES)     Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
 <210> sequence id number  
 <400> sequence id number  
 000
- 9      Use of n's or Xaa's  
    (NEW RULES)     Use of n's and/or Xaa's have been detected in the Sequence Listing.  
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response     Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11      Use of <220>     Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses  
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
 (See "Federal Register," 00/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"     Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n/Xaa     "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/650,365

DATE: 11/21/2003

TIME: 11:00:40

Input Set : A:\sequence listing.ST25.txt

Output Set: N:\CRF4\11212003\J650365.raw

3 <110> APPLICANT: Wei, Guangwen  
 4 Guo, Rongbing  
 5 Zhang, Renhuai  
 7 <120> TITLE OF INVENTION: Recombinant Super-Compound Interferon  
 9 <130> FILE REFERENCE: 792-A-PCT-US  
 11 <140> CURRENT APPLICATION NUMBER: 10/650,365  
 12 <141> CURRENT FILING DATE: 2003-08-28  
 14 <150> PRIOR APPLICATION NUMBER: PCT/CN02/00128  
 15 <151> PRIOR FILING DATE: 2002-02-28  
 17 <150> PRIOR APPLICATION NUMBER: CN 01104367.9  
 18 <151> PRIOR FILING DATE: 2001-02-28  
 20 <160> NUMBER OF SEQ ID NOS: 13  
 22 <170> SOFTWARE: PatentIn version 3.2  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 504  
 26 <212> TYPE: DNA  
 27 <213> ORGANISM: human synthesis  
 30 <220> FEATURE:  
 31 <221> NAME/KEY: CDS  
 32 <222> LOCATION: (1)..(504)  
 34 <400> SEQUENCE: 1  
 35 atg tgc gac ctg ccg cag acc cac tcc ctg ggt aac cgt cgt gct ctg 48  
 36 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu  
 37 1 5 10 15  
 39 atc ctg ctg gct cag atg cgt cgt atc tcc ccg ttc tcc tgc ctg aaa 96  
 40 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys  
 41 20 25 30  
 43 gac cgt cac gac ttc ggt ttc ccg cag gaa gaa ttc gac ggt aac cag 144  
 44 Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln  
 45 35 40 45  
 47 ttc cag aaa gct cag gct atc tcc gtt ctg cac gaa atg atc cag cag 192  
 48 Phe Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln  
 49 50 55 60  
 51 acc ttc aac ctg ttc tcc acc aaa gac tcc tcc gct gct tgg gac gaa 240  
 52 Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu  
 53 65 70 75 80  
 55 tcc ctg ctg gaa aaa ttc tac acc gaa ctg tac cag cag ctg aac gac 288  
 56 Ser Leu Leu Glu Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp  
 57 85 90 95  
 59 ctg gaa gct tgc gtt atc cag gaa gtt ggt gtt gaa gaa acc ccg ctg 336  
 60 Leu Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu  
 61 100 105 110  
 63 atg aac gtt gac tcc atc ctg gct gtt aaa aaa tac ttc cag cgt atc 384

Does Not Comply  
Corrected Diskett Ne ded

(p.1-5)

Invalid response  
please see item  
10 on error summary  
sheet

## RAW SEQUENCE LISTING

DATE: 11/21/2003

PATENT APPLICATION: US/10/650,365

TIME: 11:00:40

Input Set : A:\sequence listing.ST25.txt

Output Set: N:\CRF4\11212003\J650365.raw

64 Met Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile  
 65 115 120 125  
 67 acc ctg tac ctg acc gaa aaa aaa tac tcc ccg tgc gct igg gaa gtt 432  
 68 Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val  
 69 130 135 140  
 71 gtt cgt gct gaa atc atg cgt tcc ttc tcc ctg tcc acc aac ctg cag 480  
 72 Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln  
 73 145 150 155 160  
 75 gaa cgt ctg cgt cgt aaa gaa taa 504  
 76 Glu Arg Leu Arg Arg Lys Glu  
 77 165  
 80 <210> SEQ ID NO: 2  
 81 <211> LENGTH: 167  
 82 <212> TYPE: PRT  
 83 <213> ORGANISM: human synthesis  
 85 <400> SEQUENCE: 2  
 87 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu  
 88 1 5 10 15  
 91 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys  
 92 20 25 30  
 95 Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln  
 96 35 40 45  
 99 Phe Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln  
 100 50 55 60  
 103 Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu  
 104 65 70 75 80  
 107 Ser Leu Leu Glu Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp  
 108 85 90 95  
 111 Leu Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu  
 112 100 105 110  
 115 Met Asn Val Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Gln Arg Ile  
 116 115 120 125  
 119 Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val  
 120 130 135 140  
 123 Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln  
 124 145 150 155 160  
 127 Glu Arg Leu Arg Arg Lys Glu  
 128 165  
 131 <210> SEQ ID NO: 3  
 132 <211> LENGTH: 360  
 133 <212> TYPE: DNA  
 134 <213> ORGANISM: human synthesis  
 137 <220> FEATURE:  
 138 <221> NAME/KEY: CDS  
 139 <222> LOCATION: (1)..(360)  
 141 <400> SEQUENCE: 3  
 142 atg tgt gat tta cct caa act cat tct ctt ggt aac cgt cgc gct ctg 48  
 143 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu  
 144 1 5 10 15

-Invalid response  
 please see item  
 10 on error summary  
 sheet

-Invalid response

## RAW SEQUENCE LISTING

DATE: 11/21/2003

PATENT APPLICATION: US/10/650,365

TIME: 11:00:40

Input Set : A:\sequence listing.ST25.txt

Output Set: N:\CRF4\11212003\J650365.raw

```

146 att ctg ctg gca cag atg cgt cgt att tcc ccg ttt agc tgc ctg aaa      96
147 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys
148          20          25          30
150 gac cgt cac gac ttc ggc ttt ccg caa gaa gag ttc gat ggc aac caa      144
151 Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln
152          35          40          45
154 ttc cag aaa gct cag gca atc tct gta ctg cac gaa atg atc caa cag      192
155 Phe Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln
156          50          55          60
158 acc ttc aac ctg ttt tcc act aaa gac agc tct gct gct tgg gac gaa      240
159 Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu
160 65          70          75          80
162 agc ttg ctg gag aag ttc tac act gaa ctg tat cag cag ctg aac gac      288
163 Ser Leu Leu Glu Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp
164          85          90          95
166 ctg gaa gca tgc gta atc cag gaa gtt ggt gta gaa gag act ccg ctg      336
167 Leu Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu
168          100          105          110
170 atg aac gtc gac tct att ctg gca      360
171 Met Asn Val Asp Ser Ile Leu Ala
172          115          120
175 <210> SEQ ID NO: 4
176 <211> LENGTH: 120
177 <212> TYPE: PRT
178 <213> ORGANISM: human synthesis - SAME ERROR
180 <400> SEQUENCE: 4
182 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu
183 1          5          10          15
186 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys
187          20          25          30
190 Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln
191          35          40          45
194 Phe Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile Gln Gln
195          50          55          60
198 Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu
199 65          70          75          80
202 Ser Leu Leu Glu Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp
203          85          90          95
206 Leu Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr Pro Leu
207          100          105          110
210 Met Asn Val Asp Ser Ile Leu Ala
211          115          120
214 <210> SEQ ID NO: 5
215 <211> LENGTH: 108
216 <212> TYPE: DNA
217 <213> ORGANISM: chemical synthesis - SAME ERROR
220 <220> FEATURE:
221 <221> NAME/KEY: CDS
222 <222> LOCATION: (1)..(108)

```

## RAW SEQUENCE LISTING

DATE: 11/21/2003

PATENT APPLICATION: US/10/650,365

TIME: 11:00:40

Input Set : A:\sequence listing.ST25.txt

Output Set: N:\CRF4\11212003\J650365.raw

224 <400> SEQUENCE: 5

225 atg tgc gac ctg ccg cag acc cac tcc ctg ggt aac cgt cgt gct ctg 48

226 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu

227 1 5 10 15

229 atc ctg ctg gct cag atg cgt cgt atc tcc ccg ttc tcc tgc ctg aaa 96

230 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys

231 20 25 30

233 gac cgt cac gac 108

234 Asp Arg His Asp

235 35

238 <210> SEQ ID NO: 6

239 <211> LENGTH: 36

240 <212> TYPE: PRT

241 <213> ORGANISM: chemical synthesis - Same error

243 <400> SEQUENCE: 6

245 Met Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg Ala Leu

246 1 5 10 15

249 Ile Leu Leu Ala Gln Met Arg Arg Ile Ser Pro Phe Ser Cys Leu Lys

250 20 25 30

253 Asp Arg His Asp

254 35

257 <210> SEQ ID NO: 7

258 <211> LENGTH: 107

259 <212> TYPE: DNA

260 <213> ORGANISM: chemical synthesis - Same error

263 <220> FEATURE:

264 <221> NAME/KEY: Unsure

265 <222> LOCATION: (1)..(107)

267 <400> SEQUENCE: 7

268 ctgaaagacc gtcacgactt cggtttcccg caggagaggt tcgacggtaa ccagttccag 60

270 aagctcaggc tatctccggt ctgcacgaaa tgatccagca gaccttc 107

273 <210> SEQ ID NO: 8

274 <211> LENGTH: 103

275 <212> TYPE: DNA

276 <213> ORGANISM: chemical synthesis - Same error

279 <220> FEATURE:

280 <221> NAME/KEY: Unsure

281 <222> LOCATION: (1)..(103)

283 <400> SEQUENCE: 8

284 gctgctggta cagttcgggt tagaattttt ccagcagggg ttctgtcccaa gcagcggagg 60

286 agtcttttgt ggagaacagg ttgaaggtct gctggatcat ttc 103

289 <210> SEQ ID NO: 9

290 <211> LENGTH: 103

291 <212> TYPE: DNA

292 <213> ORGANISM: chemical synthesis - Same error

295 <220> FEATURE:

296 <221> NAME/KEY: Unsure

297 <222> LOCATION: (1)..(103)

299 <400> SEQUENCE: 9

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/650,365

DATE: 11/21/2003

TIME: 11:00:40

Input Set : A:\sequence listing.ST25.txt

Output Set: N:\CRF4\11212003\J650365.raw

300 atccctgctg gaaaaattct acaccgaact gtaccagcag ctgaacgacc tggaagcttg 60  
302 cgttatccag gaagttggtg ttgaagaaac cccgctgatg aac 103  
305 <210> SEQ ID NO: 10  
306 <211> LENGTH: 106  
307 <212> TYPE: DNA  
308 <213> ORGANISM: chemical synthesis - Same error  
311 <220> FEATURE:  
312 <221> NAME/KEY: Unsure  
313 <222> LOCATION: (1)..(106)  
315 <400> SEQUENCE: 10  
316 gaagaaaccc cgctgatgaa cgttgactcc atcctggctg ttaaaaaata ctccagcgt 60  
318 atcacctgt acctgaccga aaaaaaatac tccccgtgcg ctggg 106  
321 <210> SEQ ID NO: 11  
322 <211> LENGTH: 112  
323 <212> TYPE: DNA  
324 <213> ORGANISM: chemical synthesis - Same error  
327 <220> FEATURE:  
328 <221> NAME/KEY: Unsure  
329 <222> LOCATION: (1)..(112)  
331 <400> SEQUENCE: 11  
332 ttattcttta cgacgcagac gttcctgcag gttggtggac agggagaagg aacgcatgat 60  
334 ttcagcacga acaacttccc aagcgcacgg ggagtatttt ttttcggica gg 112  
337 <210> SEQ ID NO: 12  
338 <211> LENGTH: 31  
339 <212> TYPE: DNA  
340 <213> ORGANISM: chemical synthesis - Same error  
343 <220> FEATURE:  
344 <221> NAME/KEY: Unsure  
345 <222> LOCATION: (1)..(31)  
347 <400> SEQUENCE: 12  
348 atcgccata tgtgcgacct gccgcagacc c 31  
351 <210> SEQ ID NO: 13  
352 <211> LENGTH: 40  
353 <212> TYPE: DNA  
354 <213> ORGANISM: chemical synthesis - Same error  
357 <220> FEATURE:  
358 <221> NAME/KEY: Unsure  
359 <222> LOCATION: (1)..(40)  
361 <400> SEQUENCE: 13  
362 actgccaggc tgcagttatt ctttacgacg cagacgttcc 40

**VERIFICATION SUMMARY**

PATENT APPLICATION: **US/10/650,365**

DATE: 11/21/2003

TIME: 11:00:41

Input Set : **A:\sequence listing.ST25.txt**

Output Set: **N:\CRF4\11212003\J650365.raw**